Secure Replica Allocation in Cloud Storage Systems with Heterogeneous Vulnerabilities

Yun Tian
California State University, Fullerton
Email: ytian@fullerton.edu

Xiao Qin
Auburn University

Yafei Jia
IBM Research - China

Outlines

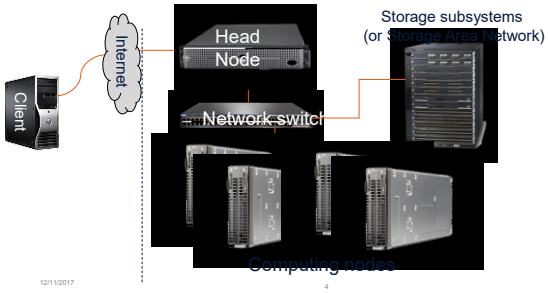
- Introduction
- Background
- SecRA– A Secure Replica Allocation Scheme
- Storage Assurance Model
- Evaluation Results
- Future Work

Cloud Storage Systems

- Large-scale data processing
- Scalability, availability, performance, security

Cloud Storage-A distributed Cluster System

- The Architecture of a Cluster



Heterogeneities in Large-Scale Distributed systems

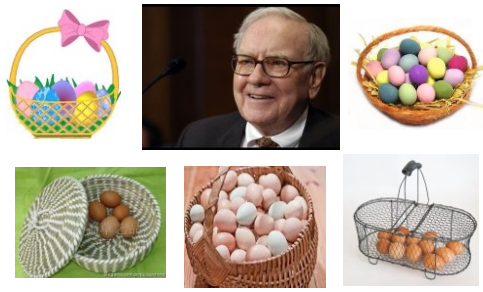
- A wide variety of:
  - Hardware (e.g., SSDs, HDDs, Tapes)
  - Software (e.g, HDFS, Lustre, PVFS)

- Heterogeneities affect performance

Can we leverage heterogeneity features to improve security for distributed systems?

Heterogeneity Influence on Storage Assurance

- Do (Not) Put Our Eggs into One Basket



Multiple server-type groups

- Diversity make sense?
  - A team with diversity make creativity
  - A system with diversity may improve security
  - Classify storage nodes of a system into different "server types" based on their different security level or strategy

Heterogeneity Impact in S-FAS(NAS '11)

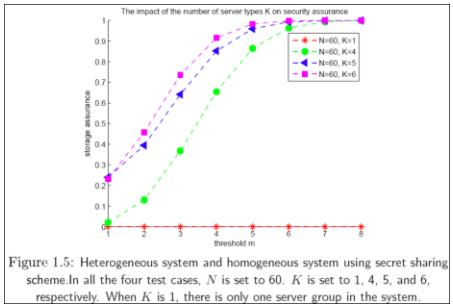


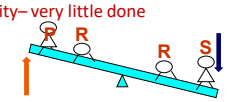
Figure 1.5: Heterogeneous system and homogeneous system using secret sharing scheme.In all the four test cases, N is set to 60. K is set to 1, 4, 5, and 6, respectively. When K is 1, there is only one server group in the system.

Data Replication to Improve Reliability

- Reliability is highly Desired
- Data Replication to enhance reliability
- There may be two cases
  - Case 1: Each file fragment has the same number *t* of replicas.
  - Case 2: File fragments have various number (i.e., *t*1, *t*2, ... , *tx*) copies of replicas.

Top Desired Properties: Security, Performance and Reliability

- Diversity—Positive or Negative?
- Diversity in a Team
  - A team with diversity is more creative
- Make Good Use of Heterogeneous Feature in Storage Nodes to Improve
  - Performance– a lot of work done
  - Security– very little done
  - Reliability– very little done



Objective

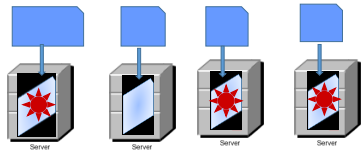
A storage solution for heterogeneous cloud storage

with data replication

considering security and performance.

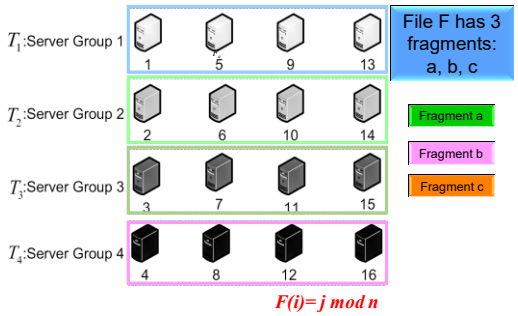
Fragmentation & Secret Sharing

- *(m, n)* Secret sharing - e.g., (2, 3)
- File Fragmentation

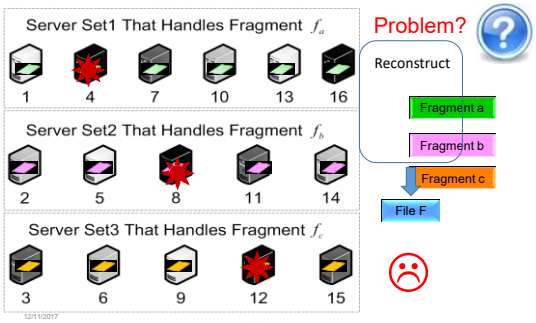


Motivation: Storage nodes in a distributed system have heterogeneous vulnerabilities.

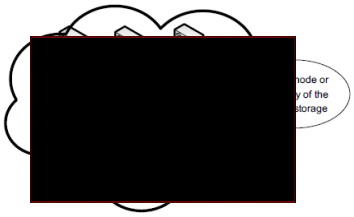
Considering Heterogeneity Issues



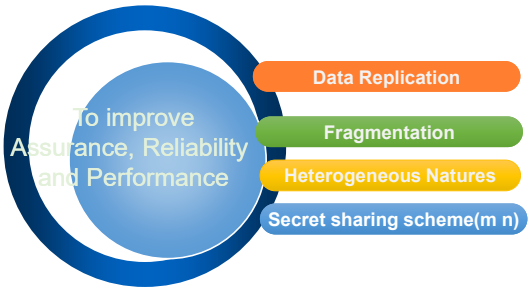
File Fragment Allocations Issues



SecRA System Model



SecRA--Outlines



SecRA--Security

- Classify storage nodes into multiple server-type groups
- Fragments of a file are assigned to storage nodes from as many different server-type groups as possible.
- Replicas of the same fragment are assigned to nodes of the same server-type

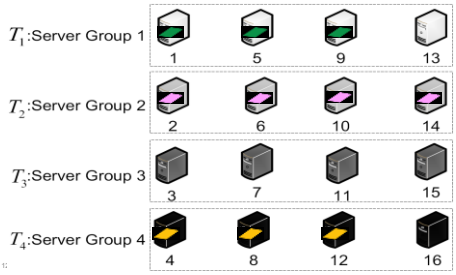
SecRA--Performance

Assigns at least one complete replica set of a file to a subsystem that is close to clients.

A Case Study for SecRA

Store fragments of a file across more server-type groups

Store replicas of a fragment across less server-type groups



Assurance Model for SecRA

- Multiplication Principle
- Conditional Probability
- Combinatorics

Assurance Model for SecRA

Static Assurance

$$SA(\alpha) = 1 - P(V)$$

$$= 1 - \sum_{j=1}^{K} \{ \frac{S_j}{N} * P(Z) * \frac{\sum_{y=m}^{S_j} [T(m,y) + T(m+1,y) + ... + T(n,y)]}{C_N^t * C_{N-t}^t * C_{N-2t}^t * ... * C_{N-(n-1)t}^t} \}$$

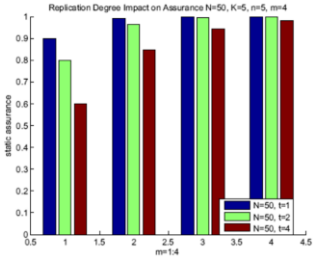
$$T(x,y) = C_x^x * C_{xt}^y - C_x^{x-1} * C_{(x-1)t}^y + ...$$

$$= \sum_{i=0}^{x-1} (-1)^i * C_x^{x-i} * C_{(x-i)t}^y$$

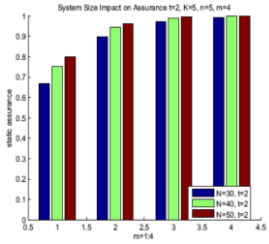
Design of Evaluation Experiments

- Replication Degree
- System Size
- Number of Fragments of a File

Preliminary Results(1)



Preliminary Results(2)



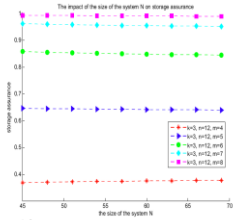


Figure 1.6: The impact of the system size N on storage assurance.

Preliminary Results(3)

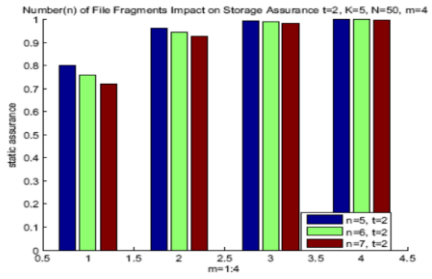


Figure 6.9: *A distributed storage system* contains 20.

Future Work

- Dynamic replica reallocation scheme
- Implement the scheme in a cloud storage system
- Evaluate SecRa by real world trace

Summary

- Introduction
- Background
- SecRA– A Secure Replica Allocation Scheme
- Storage Assurance Model
- Evaluation Results
- Future Work

Thank you!

Suggestions & Questions?